非对称周期推理循环渐进的人脸修复算法研究 *

李雅倩，张旭曜†，李岐龙

(燕山大学 电气工程学院，河北 秦皇岛 066004)

摘　要：针对生成对抗网络中修复网络无法兼顾图像的全局一致性和局部一致性，且计算负载较大的问题，在非对称 U-Net 网络架构的基础上引入渐进修复的思想。首先，提出了非对称周期特征推理模块，增加图像修复内容与周围已知像素之间的关联性，提高了修复图像的全局一致性表现。其次，提出新型的 U-Net 结构生成器网络，避免了编码器中的未知像素进入解码器，从而破坏解码器中特征的问题。最后，引入了感知损失和风格损失，进而提高了网络在主观评价下的修复效果。在人脸图像数据集上的实验表明，该文算法在主观视觉效果和客观指标上都有显著的提高。

关键词：生成对抗网络；渐进式修复；非对称周期特征推理；图像修复

中图分类号：TP391　　doi: 10.19734/j.issn.1001-3695.2022.01.0009

Asymmetric periodic inference cyclic progressive face completion algorithm research

Li Yaqian, Zhang Xuyao†, Li Qilong

(*Yanshan University*, *School of Electrical Engineering*, *Qinhuangdao Hebei* 066004, *China*)

Abstract: To address the problem that the inpainting network in generative adversarial networks cannot take into account the global consistency and local consistency of images and has a large computational load, this paper introduces the idea of incremental inpainting based on the asymmetric U-Net network architecture. Firstly, this paper proposed an asymmetric periodic feature inference module to increase the correlation between image inpainting content and surrounding known pixels, which improves the global consistency performance of restored images. Secondly, this paper presented a novel U-Net structured generator network to avoid the problem of unknown pixels in the encoder entering the decoder and thus corrupting the features in the decoder. Finally, it introduced perceptual loss and style loss, which in turn improve the inpainting effectiveness of the network under subjective evaluation. Experiments on the face image dataset show that the algorithm in this paper shows significant improvements in both subjective visual effects and objective metrics.

Key words: generative adversarial network; gradual inpainting; asymmetric periodic feature inference; image inpainting

0　引言

图像作为人们日常生活中一个重要的信息载体，当其产生破损、老化或者部分信息丢失的时候很容易使人们对信息产生误解。由于现实中的图像是在非约束条件下获得的，所以获得的图像可能会出现遮挡、污渍、破损等情况，当图像中出现上述情况的时候会明显对图像的识别、检测与分割产生负面影响。因此图像修复技术是一种被现实迫切需要的技术，特别是在如今图像识别、分割等任务已经进入日常使用环境下的今天。

传统的工匠修复照片通常首先对图像的线条或者物体轮廓进行修补，之后再对图像的颜色和纹理进行修复。Li 等人[1]首次将 GAN 应用在人脸修复领域，借鉴工匠修复图片先轮廓后纹理的思路，在鉴别器中同时引入全局和局部损失函数，生成的图片不仅在语义上有效，视觉上也比较和谐。由此图像修复技术发展出了一类将修复任务分级的修复方法。Nazeri 等人[2]提出了一种线条-颜色两步修复算法，先用一个 Patch GAN[3]进行图像的线条修复，再用另一个 Patch GAN 进行颜色修复，这样两个 GAN 串联的方式可以将修复网络分级，得到精细度更好的修复图像。Xiong 等人[4]提出应明确区分结构推断和内容修复。所提出的模型通过精确的预测边界来指导图像修复，边界修复模块能够在待修复区域推断出合理的结构，图像修复模块再根据预测边界来生成图像内容，同时将粗网络-细网络修复[5]的方法融合到了每一个 GAN 中。还有一种类似的方法是基于分层 VQ-VAE 的 multiple-solution 图像修复方法[6]，这种方法的模型将结构和纹理分开提出为结构特征的分布学习一个条件自回归网络，对于纹理生成,提出一个结构注意力模块，用来捕捉结构特征的远距离相关性，改善了结构的一致性和纹理的真实性[7]。Li 等人设计了一种线条修复模块将线条修复过程简化，从而可以使用一个 GAN 进行线条-颜色修复的方法。通过设计可视化结构重构(visual structure reconstruction, VSR)模块[8]可在生成器的初始层中进行线条修复，在生成器的主干通道中进行颜色修复。将线条修复作为一个模块嵌入到网络中可以有效的避免分级修复中多个 GAN 串联造成的计算负载大和不易收敛的问题。随后 Li 等人[9]认为线条-颜色分级不是图像修复的最优解，因此设计了周期特征推理(recurrent feature reasoning, RFR)模块。它采取修复区域分级的方法，逐步向内修复。RFR 模块可以控制每一步修复的区域，然后控制下一步修复采用上次修复区域的内容作为推理依据进行下一步的填充，通过在各层推理特征图中设计知识一致性注意力(knowledge consistent attention, KCA) [9]机制来增强各层推理的相关性。

上述工作中，大部分模型虽然获得了更好的图像修复效果，但也随之增加了计算负担，并且没有解决多个 GAN 串联不易收敛的问题。如何在不增加训练难度的同时引入分级修复的思想，针对修复任务设计出一种即插即用的特征推理

收稿日期：2022-01-02；**修回日期：**2022-02-24　**基金项目：**国家自然科学基金资助项目(62106214)；河北省自然科学基金资助项目(F2019203195)

作者简介：李雅倩(1982-)，女，湖南岳阳人，副教授，硕导，博士，主要研究方向为图像处理、机器视觉；张旭曜(1996-)，男(通信作者)，河北石家庄，硕士研究生，主要研究方向为计算机视觉(1106343910@qq.com)；李岐龙(1995-)，男，河南平顶山，硕士研究生，主要研究方向为图像修复.

模块是一种简单有效的解决思路。RFR 模块为渐进修复方法提供了一种模块化设计网络的新思路，受启发于 RFR 的设计理念，本文将 RFR 模块进行重构并结合非对称 U-Net 框架，设计了非对称周期特征推理模块，从而进一步提升模块的特征推理能力。本文的主要贡献如下：

(1)本文针对图像修复网络无法兼顾图像的全局一致性和局部一致性，且需要的计算负载较大的问题。比较现有方法后发现 RFR 模块在降低编码器部分的未知像素，提高网络的解码有良好表现。将 RFR 模块进行重构改进成非对称周期特征模块应用 U-Net 结构之中，提出了非对称周期特征推理模块。

(2)本文提出了一种渐进修复[10]的人脸修复网络。引入 PatchGAN 训练修复网络，使用门控卷积层作为 U-Net 结构的生成器的编码器中的卷积函数，U-Net 底层改为双通道空洞卷积的方式，结合感知损失和风格损失，获得了良好的人脸修复效果。

(3)实验在 CelebA 人脸数据集上采用了随机 mask 作为训练和测试数据，并且与 Edge-Connection[2]，PConv[11]，GatedConv[12]，GFP-GAN[13]和 LaMa[14]现有先进的修复模型进行对比，结果表明本文算法在提升主观效果和客观指标上的有效性。

1 算法原理

1.1 循环修复网络框架

周期特征推理模块启发于渐进修复思想以及部分卷积中的修复区域定位机制，通过对 mask 通道的更新进行每一步修复区域的定位。周期推理模块分为四个部分：区域定位(area identification)[15]特征推理(feature reasoning)[16]特征融合(feature merging)知识一致性注意力(knowledge consistent attention)，模块结构如图 1 所示。

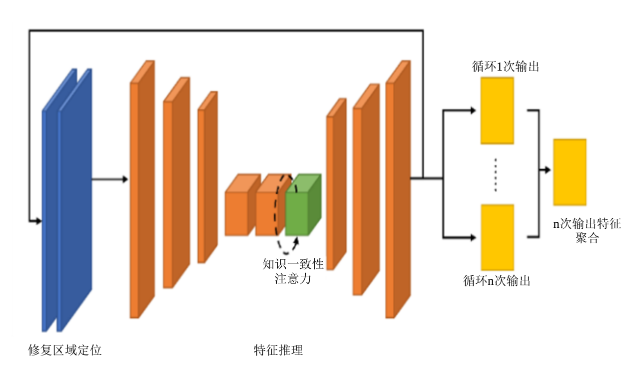


图 1 周期推理模块

Fig. 1 Periodic reasoning module

和现在流行的图像修复方法不同的是，RFR 模型中并没有使用 GAN，网络总体上是一个 CNN 结构。在本文设计非对称周期特征推理模块中，将 GAN 优良的无监督学习能力引入，设计 PatchGAN 结构，采用基于 Wasserstein 距离[17]设计的马尔可夫判别器训练嵌入了 RFR 模块的 U-Net 生成器。区域定位采用部分卷积的 mask[18]更新机制定位每一个循环体中修复区域的大小和位置。由于基于普通卷积的卷积方式不适合用于图像空洞填充，因为空间共享的卷积过滤器将所有输入像素或特征视为相同的有效特征；会导致诸如颜色差异、模糊和孔周围明显的边缘反映等视觉假象，所以本文采用门控卷积[12]作为特征推理模块的卷积核使之提高对修复内容的特征推理能力，构建一个非对称的结构。本文提出的非对称周期特征推理模块框架如图 2 所示。

采用上述非对称特征推理模块结合 U-Net 网络结构的跳跃连接[19]机制提高解码器的图像生成能力，从而提高图像的视觉表现，采用在人脸数据集上预训练的 VGG-Net 构建感知损失和风格损失可以提升生成图在图像视觉感知和风格上的表现。设计的非对称周期推理循环渐进的人脸修复网络框架如图 3 所示。其中循环渐进修复网络首先将 RFR 模块嵌入在输入层后面使得进入修复网络的编码器的时候图像已经没有未知区域，随后将不含未知区域的特征图送入 U-Net 网络中，此网络类似 RNN，特征图第一次输入到 RFR 模块里输出的结果又会第二次重新输入到 RFR 模块里。第二次则在第一次填充的特征图的基础上进一步填充或预测出更合理的值。这样反复几个循环后才会进入到下一步的特征融合环节。这种循环机制带来的好处是能反复利用 RFR 模块中的特征，使模型达到轻便化的同时预测出更合理的特征参数，进一步提升模块特征推理能力。为了提高图像在高层特征图中提取语义的能力，将 U-Net 的底层设计为并联空洞卷积结构。该算法使用门控卷积层作为 U-Net 结构生成器中编码器的卷积函数，U-Net 底层改为双通道空洞卷积的方式以防止图像空间结构在生成器的底层被破坏。

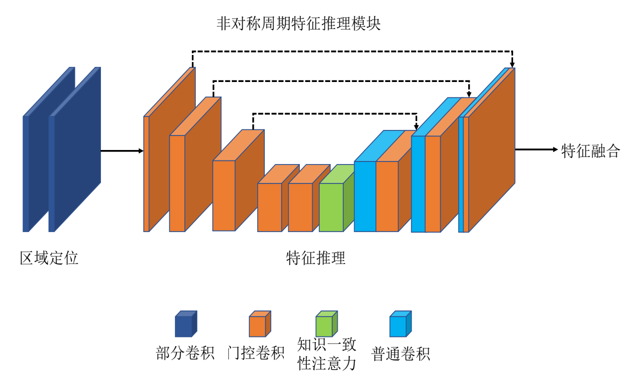


图 2 非对称周期特征推理

Fig. 2 Asymmetric periodic feature inference

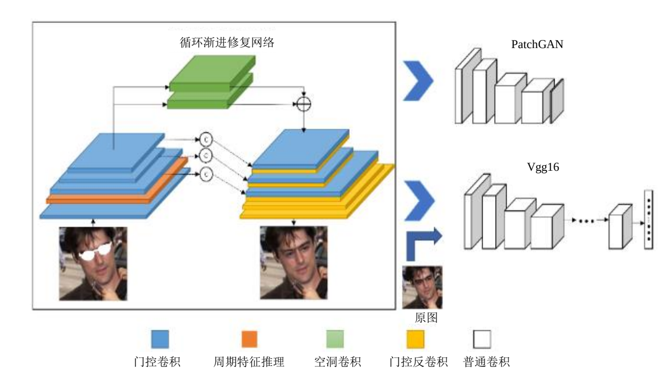


图 3 非对称周期推理循环渐进的人脸修复模型

Fig. 3 Asymmetric cycle inference for cyclic progressive face restoration model

1.2 马尔可夫判别器

普通的 GAN 思想是采用一个判别器来判定输入是来自真实数据分布 p_r 还是生成数据分布 p_g。与普通 GAN 的判别器不同，马尔可夫判别器设计为全卷积结构，输出为一个 $N \times N$ 的矩阵 X，其中矩阵 X 的元素 $X_{i,j}$ 代表在位置 (i, j) 处对应到原图一个 $M \times M$ 大小的感受野内图像块的真实度，最后将矩阵所有的值取平均得到图像的真实度。网络训练采用基于 Wasserstein 距离的衡量指标，保证可以收敛。

采用马尔可夫判别器相对于普通判别器来说有以下优点：

(1)马尔可夫判别器可以关注图像块的修复结果，能够显著提升图像的细节表现。

(2)通过关注输入的不同区域，可以通过学习考虑到图中不同区域对于识别该图是否属于真实图像的贡献度，从而可以有针对性的考虑贡献大的区域。例如在针对人脸进行修复

的时候，判别器重点关注人脸区域，减少对背景区域的关注。

为了提高 PatchGAN 的收敛速度，同时提高修复的多样性，参考了 Miyato 等人的研究，在 PatchGAN 中引入谱范数正则化。WGAN 能够收敛所需要的条件是网络满足 Lipschitz 约束[20]。为了满足这个条件，原始 WGAN 采用了梯度裁剪的方式人为的将所有卷积层的权重约束到一个范围内，最后造成了梯度分布不均匀的问题。随后提出的 WGAN-G[21]采用梯度惩罚的正则化方式缓解了这个问题。后来 Miyato 等人提出了采用谱范数正则化的方式，进而在不对梯度加以干扰的条件下使得网络满足 1-Lipschitz 约束。本文所采用的马尔可夫判别器中采用谱范数正则化进行训练，从而提高判别器的效果。

1.3　感知损失与风格损失

针对人脸修复任务，本文引进了感知损失与风格损失[22]。本文算法采用在人脸数据集上预训练的人脸识别网络作为监督，对修复图和原图进行感知差别的计算和风格保持的计算。其中修复网络的输出和原图的尺寸为 256×256，而 VGG16 网络的输入为 224×224 的图像，因此在训练时将修复网络的输出图与原图在计算损失时尺寸变为 224×224。所用的 VGG16 网络结构如图 4 所示。

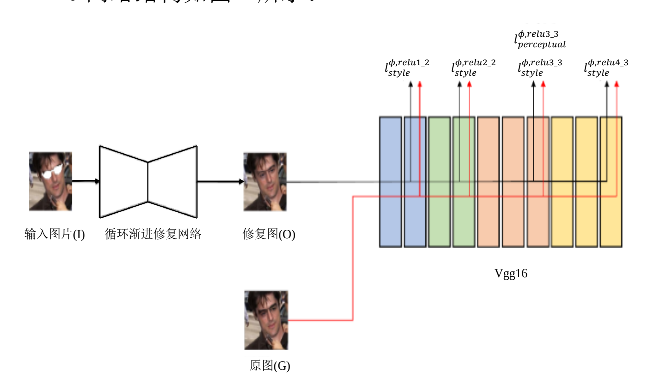


图 4　感知损失与风格损失计算

Fig. 4　Perceived loss and style loss calculation

感知损失的计算在 VGG16 网络的 relu3_3 层计算，计算方式如式(1)所示。

$$l_{\text{perceptual}}^{\phi,j}=\frac{1}{C_jH_jW_j}\left\|\phi_j(0)-\phi_j(g)\right\|_2^2, \tag{1}$$

其中，ϕ 代表 VGG 网络，j 代表使用 VGG 网络的第 j 层的特征图作为输出，C_j，H_j，W_j 分别代表 VGG 网络的第 j 层输出的特征图的通道数、高度和宽度。

采用 VGG16 作为监督网络，识别输入两个图在特征图层面的像素级差别。与直接在图像上进行像素级差别计算相比，特征图可以表示图像的高层语义特征，因此在特征图上进行像素级的计算可以表示两张图之间的感知差别。

风格损失计算方式如式(2)及(3)所示。

$$G_j^{\phi}(x)_{c,c'}=\frac{1}{C_jH_jW_j}\sum_{h=1}^{H_j}\sum_{w=1}^{W_j}\phi_j(x)_{h,w,c}\phi_j(x)_{h,w,c'}, \tag{2}$$

$$l_{\text{style}}^{\phi,j}(o,g)=\left\|G_j^{\phi}(o)-G_j^{\phi}(g)\right\|_F^2, \tag{3}$$

其中，ϕ 代表 VGG 网络，j 代表使用 VGG 网络的第 j 层的特征图作为输出，C_j，H_j，W_j 分别代表 VGG 网络的第 j 层输出的特征图的通道数、高度和宽度，G 代表计算格拉姆矩阵。

1.4　总体损失

训练网络总共包含四部分的损失，即 PatchGAN 损失、感知损失、风格损失及 L1 重构损失。定义网络的输入图片为 $I_{in}=I_{gt}\odot M$，式中 M 是掩码图像，在 M 中像素点为 0 的区域为未知区域，像素点为 1 的区域为已知区域，输入的破损图像是由原始图像和掩码图像进行逐像素的叠加得到。输入图像经过循环渐进人脸修复网络的修复，输出为 $O=I_{gt}\odot M+G(I,M)\odot(1-M)$，式中 G 为网络的生成器，输入图片未被遮挡的原图为 I_{gt}，网络的总损失函数如下式(4)所示。

$$L=\lambda_{adv}L_{adv}+\lambda_{per}L_{per}+\lambda_{style}L_{style}+\lambda_{l1}L_{l1}, \tag{4}$$

其中，L_{adv} 代表生成对抗网络损失[23]，L_{per} 代表感知损失，L_{style} 代表风格损失，L_{l1} 代表 L1 重构损失。

本文网络的优化方法使用 Adam 优化器，针对总体损失进行优化。首先对网络进行前向计算，得到输出的修复图像，随后根据式(4)计算网络的损失，采用 Adam 对损失进行优化。所计算的权重如下所示。

$$w=w-\frac{\eta m_t}{\sqrt{n_t}+\varepsilon}, \tag{5}$$

上式中的一阶动量 m_t 和二阶动量 n_t 计算由式(6)和(7)给出：

$$m_t=\frac{m_t}{1-\gamma_1^t}=\frac{\gamma_1^t m_{t-1}+(1-\gamma_1^t)}{1-\gamma_1^t}\frac{\partial L}{\partial w}, \tag{6}$$

$$n_t=\frac{n_t}{1-\gamma_2^t}=\frac{\gamma_2^t m_{t-1}+(1-\gamma_2^t)}{1-\gamma_2^t}\frac{\partial L}{\partial w}, \tag{7}$$

其中，γ_1^t、γ_2^t 和 ε 都是优化算法的参数，L 是网络的总体损失。梯度反传过程首先通过网络的损失对权重的梯度计算出一阶动量和二阶动量的无偏估计，代入式(5)即可更新权重。

2　修复实验结果

2.1　深度学习环境配置

本算法的实验代码采用 PyTorch 框架实现。Pytorch 是 Facebook 公司开发的动态计算流图的深度学习框架。PyTorch 框架提供了根据计算图求各节点导数的功能，可以动态的建立计算流图，能够更加方便的实现梯度的反向传播。同时集成了大量的深度学习常用函数，方便网络搭建和实验。

本文算法的实验平台的操作系统是 64 位的 Linux 内核系统，版本是 Ubuntu 18.04LTS，PyTorch 版本为 1.2.0，OpenCV 为 4.4.0.46，numpy 为 1.16.5，scipy 为 1.1.0，scikit-image 为 0.13.1。显卡型号为 Nvidia 1080Ti 8G。

2.2　数据集介绍与超参数配置

CelebA[24]是香港中文大学收集制作的大尺度人脸数据集，数据集中一共包含 202599 张人脸图片。由于未处理的数据集中图像尺寸不一，人脸位置不一，不方便进行学习，因此首先采用作者给出的人脸 bbox 回归框标注文件进行预处理，得到 202599 张对齐后的图片，随后进行尺寸变换为 256×256 的图片，然后在图片中随机抽取 1000 张作为测试集，其余图片均为训练集。

本文的实验使用 Adam 优化器，训练时学习率为 1×10^{-4}，模型调整优化阶段(fine-tuning)学习率为 1×10^{-5}，模型的输入批次为 6，训练 450000 次循环。超参数设置为 $\lambda_{adv}=\lambda_{per}=0.05$，$\lambda_{style}=120$，$\lambda_{l_1}=1$。

2.3　实验结果与评价指标

本文算法选取的对比模型都是在人脸数据上有优秀表现的模型。横向对比的模型有 Edge-Connection[2]，PConv[11]，GatedConv[12]，GFP-GAN[13]和 LaMa[14]。本文采用 PConv 中所提出的不规则遮挡(irregular mask)数据集作为训练和测试的 mask，该数据集提供了不同遮挡比例和不同形状的随机遮挡，可以模拟现实世界中遮挡的随机性，采用该数据集作为测试 mask 对于横向对比而言更有客观性，不同比率的 mask 遮挡如图 5 所示，待修复图片如图 6 所示。

各算法在人脸数据集上的主观表现如图 7 所示。观察可知，PConv、GatedConv、Edge-Connection、GFP-GAN、LaMa 和本文算法均对还原图像的全局语义有较好的效果，但本文算法在还原图像的细节结构有更加突出的表现，前五种算法无法准确的全部还原图像中的一些细节结构。


(a)0.01-0.1的 mask


(b)0.1-0.2的 mask


(c)0.2-0.3的 mask


(d)0.3-0.4的 mask


(e)0.4-0.5的 mask


(f)0.5-0.6的 mask

图 5　不同 mask 比率示意图

Fig. 5　Diagram of different mask ratios


(a)0.01-0.1 mask下待修复图像


(b)0.1-0.2 mask下待修复图像


(c)0.2-0.3 mask下待修复图像


(d)0.3-0.4 mask下待修复图像


(e)0.4-0.5 mask下待修复图像


(f)0.5-0.6 mask下待修复图像

图 6　不同比例随机遮挡下的待修复图像

Fig. 6　Images to be restored at different scales with random masking

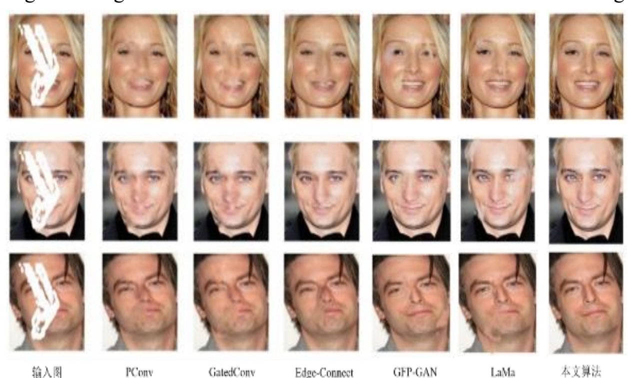


图 7　不同模型主观修复对比

Fig. 7　Comparison of subjective inpainting of different models

为了探索本文算法所提出的非对称周期特征推理模块和采用 PatchGAN 训练的网络框架的有效性，设计了与 RFR 进行对比的实验，对比结果如表 1 所示。

表 1　mask 在(0.1,0.2]时与 RFR 对比结果

Tab. 1　Mask compared the results with the original paper at (0.1,0.2]

模型	PSNR	SSIM	L1
RFR-Net	33.56	0.981	0.0075
本文算法	**33.61**	**0.983**	**0.0071**

不规则遮挡数据集将不同比例的 mask 进行了分类，其中每一个比例含有 2000 张随机 mask 图像，为了测试模拟本文算法在不同 mask 比例下的修复效果，选取 (0.1,0.2]、(0.3,0.4]以及(0.5,0.6]的三个不同比例的 mask 数据集作为测试。不同的 mask 比例下本文算法与其他算法的客观指标横向对比结果如表 2~4 所示。在 mask 比率为(0.1,0.2]时不同模型的修复效果对比如表 2 所示。

表 2　mask 在(0.1,0.2]时不同模型的指标

Tab. 2　Mask at (0.1,0.2] different model indicators

模型	PSNR	SSIM	L1
PConv	32.77	0.977	0.0083
GatedConv	32.56	0.973	0.0088
Edge-Connection	32.48	0.975	0.0088
GFP-GAN	32.86	0.979	0.0082
LaMa	33.17	0.980	0.0078
本文算法	**33.61**	**0.983**	**0.0071**

在 mask 比率为(0.3,0.4]时不同模型的修复效果对比如表 3 所示。

表 3　mask 在(0.3,0.4]时不同模型的指标

Tab. 3　Mask at (0.3,0.4] different model indicators

模型	PSNR	SSIM	L1
PConv	26.94	0.922	0.0236
GatedConv	26.72	0.914	0.0245
Edge-Connection	26.62	0.915	0.0247
GFP-GAN	27.04	0.920	0.0235
LaMa	27.15	0.926	0.0231
本文算法	**27.78**	**0.935**	**0.0211**

在 mask 比率为(0.5,0.6]时不同模型的修复效果对比如表 4 所示。

表 4　mask 在(0.5,0.6]时不同模型的指标

Tab. 4　Mask at (0.5,0.6] different model indicators

模型	PSNR	SSIM	L1
PConv	22.14	0.791	0.0524
GatedConv	21.47	0.767	0.0561
Edge-Connection	21.49	0.759	0.0572
GFP-GAN	21.62	0.792	0.0506
LaMa	22.14	0.797	0.0498
本文算法	**22.91**	**0.823**	**0.0468**

表 2~4 展现了在不同 mask 遮挡比例下本文算法的客观有效性。客观指标对比采用三种不同的指标，分别是峰值信噪比(PSNR)、结构相似性(SSIM)以及 L1 误差。其中 PSNR 与 SSIM 指标是越大越好，L1 误差是越小越好，L1 误差是修复图像与原图之间的 L1 范数。在上述三种客观指标的对比中，本文算法在不同数据集上的实验结果显示出了本文算法的优越性。

实验表明本章算法相比原算法有一定程度的提高。在实验中发现非周期特征推理的迭代次数对修复结果有影响作用，通过设计不同迭代次数的实验得到在迭代次数为 6 次的时候网络的修复性能表现最好。实验如表 5 所示。

表 5　周期推理模块迭代次数的影响

Tab. 5　Influence of iteration number of periodic reasoning module

迭代次数	4	5	6	7	8
PSNR	27.32	27.70	**27.78**	27.72	27.56
SSIM	0.914	0.928	**0.935**	0.931	0.917

上述的实验证明了本文创新的有效性，图 8 展示了本文算法在不同比例遮挡下的修复表现，图 9 展示了本文算法在同一图像上不同比例遮挡下的表现。从主观效果上来看，本文算法在较大范围的任意遮挡上获得了比较好的修复效果。

3　结束语

针对修复网络无法兼顾图像的全局一致性和局部一致性，且需要的计算负载较大的问题，本文提出了非对称周期特征推理模块，提高了图像修复内容与周围已知像素之间的关联性。首先引入渐进修复的思想，将一种渐进修复模块(RFR)结合非对称网络的构架，对渐进修复模块进行重构；其次采用

加入了非对称 RFR 模块的 U-Net 框架，引入 PatchGAN 训练修复网络，使用门控卷积层作为 U-Net 结构的生成器的编码器中的卷积函数，U-Net 底层改为双通道空洞卷积的方式防止图像空间结构在生成器的底层被破坏；最后结合感知损失和风格损失，获得了良好的人脸修复效果。实验结果表明，本文算法分别相比于现有一些算法，在信噪比(PSRN)、结构相似性(SSIM)和 L1 误差各项客观指标以及主观视觉上都有显著的提高。



(a)0.01-0.1mask下修复图像

(b)0.1-0.2mask下修复图像

(c)0.2-0.3mask下修复图像



(d)0.3-0.4mask下修复图像



(e)0.4-0.5mask下修复图像



(f)0.5-0.6mask下修复图像

图 8 不同 mask 比例下的修复效果

Fig. 8 Repair effect under different mask ratio

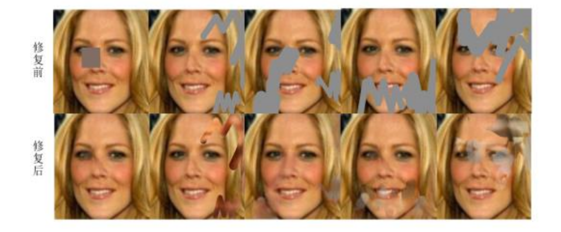


图 9 同一图像不同 mask 比例下的修复效果

Fig. 9 The repair effect of the same image with different mask proportions

参考文献：

[1] Li Yijun, Liu Sifei, Yang Jimei, *et al.* Generative face completion [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2017: 3911-3919.

[2] Nazeri K, Ng E, Joseph T, *et al.* Edgeconnect: Generative image inpainting with adversarial edge learning [J]. arXiv preprint arXiv: 1901. 00212, 2019.

[3] 杜雪梅. 基于 GAN 的图像补全算法研究 [D]. 成都: 电子科技大学, 2021. (Du Xuemei. Research on image completion algorithm based on GAN [D]. ChengDu: University of Electronic Science and Technology of China, 2021.)

[4] Xiong Wei, Yu Jiahui, Lin Zhe, *et al.* Foreground-aware image inpainting [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 5840-5848.

[5] Moskalenko A, Erofeev M, Vatolin D. Deep Two-Stage High-Resolution Image Inpainting [J]. CEUR Workshop Proceedings, 2020, 2744.

[6] Peng Jialun, Liu Dong, Xu Songcen, *et al.* Generating Diverse Structure for Image Inpainting With Hierarchical VQ-VAE [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 10775-10784.

[7] Liao Liang, Xiao Jing, Wang Zheng, *et al.* Image Inpainting Guided by Coherence Priors of Semantics and Textures [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 6539-6548.

[8] Li Jingyuan, He Fengxiang, Zhang Lefei, *et al.* Progressive reconstruction of visual structure for image inpainting [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 5962-5971.

[9] Li Jingyuan, Wang Ning, Zhang Lefei, *et al.* Recurrent Feature Reasoning for Image Inpainting [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 7760-7768.

[10] Wang Tengfei, Ouyang Hao, Chen Qifeng. Image Inpainting with External-internal Learning and Monochromic Bottleneck [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 5120-5129.

[11] Liu Guilin, Reda F A, Shih K J, *et al.* Image inpainting for irregular holes using partial convolutions [C]// Proceedings of the European Conference on Computer Vision (ECCV) . Springer, 2018: 85-100.

[12] Yu Jiahui, Lin Zhe, Yang Jimei, *et al.* Free-form image inpainting with gated convolution [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 4471-4480.

[13] Wang Xintao, Li Yu, Zhang Honglun, *et al.* Towards real-world blind face restoration with generative facial prior [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 9168-9178.

[14] Suvorov R, Logacheva E, Mashikhin A, *et al.* Resolution-robust Large Mask Inpainting with Fourier Convolutions [C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2022: 2149-2159.

[15] Gong Yunpeng. A general multi-modal data learning method for Person Re-identification [J]. arXiv preprint arXiv: 2101. 08533, 2021.

[16] Ai Mingxi, Xie Yongfang, Tang Zhaohui, *et al.* Deep learning feature-based setpoint generation and optimal control for flotation processes [J]. Information Sciences, 2021, 578: 644-658.

[17] Jam J, Kendrick C, Drouard V, *et al.* Symmetric Skip Connection Wasserstein GAN for High-resolution Facial Image Inpainting [C]// 16th International Conference on Computer Vision Theory and Applications. 2021.

[18] 曾文雯, 杨阳, 钟小品. 基于改进 Mask R-CNN 的在架图书书脊图像实例分割方法 [J]. 计算机应用研究, 2021, 38 (11): 3456-3459+3505. (Zeng Wenwen, Yang Yang, Zong Xiaopin. An improved Mask R-CNN based method for segmentation of on-shelf book spine image instances [J]. Computer Applications Research, 2021, 38 (11): 3456-3459+3505.)

[19] 李大海, 王榆锋, 王振东. 面向遥感图像云分割问题的新型 U-Net 模型 [J]. 计算机应用研究, 2021, 38 (11): 3506-3509+3516. (Li Dahai, Wang Yufeng, Wang Zhendong. A novel U-Net model for remote sensing image cloud segmentation problem [J]. Computer Application Research, 2021, 38 (11): 3506-3509+3516.)

[20] He Zhiyu, He Jianping, Chen Cailian, *et al.* CPCA: A chebyshev proxy

and consensus based algorithm for general distributed optimization [C]// 2020 American Control Conference (ACC) . IEEE, 2020: 94-99.

[21] Wei Xiang, Gong Boqing, Liu Zixia, *et al.* Improving the improved training of wasserstein gans: A consistency term and its dual effect [J]. arXiv preprint arXiv: 1803. 01541, 2018.

[22] Cai Shaoyu, Zhu Kening, Ban Y, *et al.* Visual-Tactile Cross-Modal Data Generation using Residue-Fusion GAN with Feature-Matching and Perceptual Losses [J]. IEEE Robotics and Automation Letters, 2021, 6 (4): 7525-7532.

[23] Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial networks [J]. Communications of the ACM, 2020, 63 (11): 139-144.

[24] Liu Ziwei, Luo Ping, Wang Xiaogong, *et al.* Deep learning face attributes in the wild [C]// Proceedings of the IEEE international conference on computer vision. IEEE, 2015: 3730-3738.